LOAN DOCUMENT

PHOTOGRAPH THIS SHEET

DTIC ACCESSION NUMBER

LEVEL

0

INVENTORY

AFRL-ML-TY-TP-2000-4527

DOCUMENT IDENTIFICATION

2000

DISTRIBUTION STATEMENT A
Approved for Public Release
Distribution Unlimited

DISTRIBUTION STATEMENT

H A N D L E W I T H C A R E

ACCESSION FOR

NTIS	GRA&I	☑
DTIC	TRAC	□
UNANNOUNCED		□
JUSTIFICATION		

BY

DISTRIBUTION/

AVAILABILITY CODES

DISTRIBUTION	AVAILABILITY AND/OR SPECIAL	
A-1		

DISTRIBUTION STAMP

DATE ACCESSIONED

DATE RETURNED

20000601 070

DATE RECEIVED IN DTIC

REGISTERED OR CERTIFIED NUMBER

PHOTOGRAPH THIS SHEET AND RETURN TO DTIC-FDAC

DOCUMENT PROCESSING SHEET

PREVIOUS EDITIONS MAY BE USED UNTIL STOCK IS EXHAUSTED.

LOAN DOCUMENT

AFRL-ML-TY-TP-2000-4527

INFRARED SPECTRAL CLASSIFICATION WITH ARTIFICIAL NEURAL NETWORKS AND CLASSICAL PATTERN RECOGNITION

HOWARD T. MAYFIELD
AFRL/MLQL
139 BARNES DRIVE, STE 2
TYNDALL AFB FL 32403-5323

DELYLE EASTWOOD
AIR FORCE INSTITUTE OF TECHNOLOGY
DEPARTMENT OF ENGINEERING PHYSICS
2950 P STREET
WPAFB OH 45433

LARRY W. BURGGRAF
UNIVERSITY OF SOUTH CAROLINA
DEPARTMENT OF CHEMISTRY AND BIOCHEMISTRY
712 S. MAIN STREET
COLUMBIA SC 29208

MATERIALS & MANUFACTURING DIRECTORATE
AIR FORCE RESEARCH LABORATORY
AIR EXPEDITIONARY FORCES TECHNOLOGIES DIVISION
139 BARNES DRIVE, STE 2
TYNDALL AFB FL 32403-5323

MLQL/Dr. Mayfield/File

TECHNICAL PUBLICATION / PRESENTATION CONTROL RECORD

1. TITLE Infrared Spectral Classification with Artificial Neural Networks and Classical Pattern Recognition

2. MATERIAL IS FOR
□ TECHNICAL REPORT
□ TECHNICAL PAPER
□ SPECIAL REPORT
□ PUBLICATION IN A JOURNAL
☑ ORAL PRESENTATION
☑ PROCEEDING
☑ ABSTRACT
□ OTHER (Specify)

3. AUTHOR(S) (LAST NAME, FIRST NAME, MI, RANK - LEAD AUTHOR FIRST)
Mayfield, Howard T.; Eastwood, DeLyle; and Burggraf, Larry W.

4. NAME OF JOURNAL OR DETAILS *(Date and Place)* OR ORAL PRESENTATION *(Indicate if Foreign)*
Publication in Proceedings of SPIE Vol. 4036 Chemical and Biological Sensing, Patrick J. Gardner, editor, expected to be published in September 2000. Also, oral proceedings at AeroSense meeting, 24-28 April 2000, Orlando, FL

5. PROJECT NO. (JON #) RAPIT80H

6. CONTRACT NO.

7. PROGRAM ELEMENT

8. WORK UNIT NO.

9. AUTHOR/CONTRACT MONITOR *(Name / Office Symbol / Ext)* Howard T. Mayfield AFRL/MLQL

10. CONTRACTOR

11. SECURITY CLASSIFICATION ☑ UNCLASSIFIED □ OTHER (Specify) □ DESTRUCTION NOTICE □ SBIR REPORT □ PHASE I □ PHASE II

12. RELEASE FOR PUBLICATION REQUIRED FROM OTHER AGENCY □ YES ☑ NO *If yes, attach copy of release*

13. JOINT PUBLICATION WITH OTHER GOVERNMENT ORGANIZATION □ YES ☑ NO *If yes, attach copy of other organization's release*

14. RELEASE FOR USE OF COPYRIGHTED MATERIAL ON MS PAGE IS REQUIRED *(attach copy of release)*

15. SPECIAL DISTRIBUTION LIST □ YES *(Include List)* □ NO

16. TECH REPORT NO. AFRL-ML-TY-TP-2000-4527

17. LAST PUBLICATION FOR WORK UNIT □ YES □ NO □ INTERIM □ FINAL

18. DISTRIBUTION STATEMENT (AFI 61-204) *(Select distribution statement from below)*

☑ A: Approved for public release, distribution unlimited

□ B: Distribution authorized to US Government agencies only (reason)(date of determination). Other requests for this document shall be referred to (controlling DoD office).

□ C: Distribution authorized to US Government agencies and their contractors (reason)(date of determination). Other requests for this document shall be referred to (controlling DoD office).

□ D: Distribution authorized to Department of Defense and US DoD contractors only (reason)(date of determination). Other requests for this document shall be referred to controlling DoD office).

□ E: Distribution authorized to DoD components only (reason)(date of determination). Other requests for this document shall be referred to controlling DoD office).

□ F: Further dissemination only as directed by (controlling office)(date of determination) or DoD higher authority.

□ X: Export Control: Distribution authorized to US Government agencies and private individuals or enterprises eligible to obtain export-controlled technical data in accordance w/DoDD 5230.25 (date of determination). Controlling DoD office is (insert).

19. REASON *(Check one)*
□ ADMINISTRATIVE OR OPERATIONAL USE
□ CONTRACTOR PERFORMANCE EVALUATION
□ CRITICAL TECHNOLOGY
□ DIRECT MILITARY SUPPORT
□ TEST & EVALUATION
□ FOREIGN GOVERNMENT INFORMATION
□ PREMATURE DISSEMINATION
□ PROPRIETARY INFORMATION
□ SOFTWARE DOCUMENTATION
□ SPECIFIC AUTHORITY

20. FINAL EDITING AND PROCESSING *

DATE IN	TO	DATE OUT	COMMENT	SIGNATURE
28 APR 00	PROJ MGR	28 Apr 00	Approved	Howard T. Mayfield
28APR00	BRANCH CHIEF	28APR00	None	JR Stauffer
	MLQ	5/8/00		
—	STINFO	—	—	
	FDPO			
10 May	PA	10 May 00	PA Case # 00-038	
10 May	DTIC	10 May	Forwarded to DTIC As a Tech. Paper	

* I HAVE REVIEWED THE ATTACHED MATERIAL AND HAVE DETERMINED THAT IT IS UNCLASSIFIED, TECHNICALLY ACCURATE AND SUITABLE FOR RELEASE

NOTICES

USING GOVERNMENT DRAWINGS, SPECIFICATIONS, OR OTHER DATA INCLUDED IN THIS DOCUMENT FOR ANY PURPOSE OTHER THAN GOVERNMENT PROCUREMENT DOES NOT IN ANY WAY OBLIGATE THE US GOVERNMENT. THE FACT THAT THE GOVERNMENT FORMULATED OR SUPPLIED THE DRAWINGS, SPECIFICATIONS, OR OTHER DATA DOES NOT LICENSE THE HOLDER OR ANY OTHER PERSON OR CORPORATION; OR CONVEY ANY RIGHTS OR PERMISSION TO MANUFACTURE, USE, OR SELL ANY PATENTED INVENTION THAT MAY RELATE TO THEM.

THIS TECHNICAL REPORT HAS BEEN REVIEWED AND IS APPROVED FOR PUBLICATION.

Timothy G. Wiley

TIMOTHY G. WILEY, Lt Col, USAF, BSC
Program Manager

Thomas B. Stauffer

THOMAS B. STAUFFER
Chief, Weapons Systems Logistics Branch

RANDY L. GROSS, Col, USAF, BSC
Chief, Air Expeditionary Forces Technologies Division

Infrared Spectral Classification with Artificial Neural Networks and Classical Pattern Recognition

Howard T. Mayfield[a], DeLyle Eastwood[b,c], and Larry W. Burggraf[b]

[a]Air Force Research Laboratory, Air Expeditionary Forces Technology Division,
139 Barnes Drive, Suite 2, Tyndall AFB, FL 32403
[b]Air Force Institute of Technology, Department of Engineering Physics,
2950 P Street, Wright-Patterson AFB, OH 45433
[c]Current Address: University of South Carolina, Department of Chemistry and
Biochemistry, 712 South Main Street, Columbia, SC 29208

Abstract

Infrared spectroscopy is an important technique for measuring airborne chemicals, for pollution monitoring and to warn of toxic compound releases. Infrared spectroscopy provides both detection and identification of airborne components. Computer-assisted classification tools, including pattern recognition and artificial neural network techniques, have been applied to a collection of infrared spectra of organophosphorus compounds, and these have successfully discriminated commercial pesticide compounds from military nerve agents, precursors, and hydrolysis products. Infrared spectra for previous tests came from a commercial infrared library, with permission, from military laboratories, and from defense contractors. In order to further test such classification tools, additional infrared spectra from the NIST gas-phase infrared library were added to the data set. These additional spectra probed the tendency of the trained classifiers to misidentify unrelated spectra into the trained classes.

Infrared spectra used in this effort were gathered from a variety of sources. Different instrument operators collected them at a number of locations, in a variety of spectral data collection designs, and they were delivered in a variety of digital formats. The spectra were treated mathematically to remove artifacts from their collection. Preprocessing techniques used included Fisher weighting and principal component analysis. Classifications were made using the k-nearest neighbor classifier, feed forward neural networks, trained with a variety of techniques, and radial basis function networks. The results from these classification techniques will be reported and compared.

Keywords

Infrared Spectroscopy Chemometrics Classification Pattern Recognition Artificial Neural Networks Radial Basis Function Networks Organophosphorus Compounds Pesticides

Introduction

Infrared spectroscopy is a valuable technique for providing structural information, compound identification, and quantitative information for a wide variety of organic and inorganic chemicals and mixtures. The absorption of infrared radiation is controlled by vibrational energy levels within molecules, thus providing the structural information associated with the technique. Molecular vibrations that alter a molecule's dipole moment give rise to vibrational energy levels. Infrared spectrometers are widely available in laboratories and as ruggedized industrial instruments. Field portable infrared spectrometers based on dispersive optics or on Fourier transform principles are becoming available for conducting remote environmental, forensic, and industrial investigations.[1-3]

International efforts to control the proliferation of chemical weapons have resulted in the adoption of the Chemical Warfare Convention (CWC) by the international community. This convention specifically bans production and stockpiling of a number of chemicals that have been used or proposed as chemical warfare agents due to exceptional toxicity toward humans. The convention provides for the possibility of on-site inspection of military and industrial sites to ensure banned materials are not being stockpiled or produced. Inspectors carrying out these duties will need portable chemical analysis equipment that can identify banned materials present in process, product, and waste samples.[4]

Several of the banned chemicals in the CWC are organophosphorus neurotoxins that are closely related to organophosphorus pesticides. Both the banned neurotoxins and the commercial organophosphorus pesticides attack their victim's nervous system by attacking the enzyme acetylcholinesterase. The loss of this enzyme interferes with the regulation of the neurotransmitter, acetylcholine, producing muscle spasms, paralysis, etc.. The organophosphorus pesticides remain valuable in commercial pesticide applications because of their rapid and easy hydrolysis in the environment, which prevents buildup of active pesticide and residues in the environment.[5]

Infrared spectroscopy is now a highly mature technique that ensures highly reproducible spectra even between laboratories and among a variety of instruments. This allows the spectra of pure substances to be conveniently collected in libraries to support qualitative and quantitative analyses. Large infrared spectrum collections have been useful in supporting computerized spectral classification experiments based on statistical pattern recognition techniques and artificial neural networks. The reproducibility of infrared spectra between instruments and laboratories has also facilitated sharing spectral results between laboratories. The need to facilitate the exchange of spectral information of all types between workers prompted the development of the JCAMP format for exchanging digital spectral data. JCAMP files are written in an ASCII text form with labeled sections to transmit spectral information, machine conditions, etc. in a format that can be accepted by any computer using the ASCII code for character-based information.[6]

Jus and Isenhour attempted to classify infrared spectra in early efforts to utilize the linear learning machine (LLM) technique to apply pattern recognition techniques to chemical data.[7] Numerous other discrete category classification techniques have been applied to infrared pattern recognition experiments, including *k*-nearest neighbor (KNN) method and SIMCA (Soft Isostructural Modeling for Class Analysis), and various forms of discriminant analysis. Principal component analysis (PCA) is commonly used to simplify and pretreat infrared spectra.[8,9] Single component samples and two- or three- component mixtures can usually be analyzed quantitatively with simple Beer's Law calculations. Quantitative analysis of more complicated mixtures may require simple linear algebra calculations based on Beer's Law, or more complicated chemometric deconvolution procedures, such as classical least squares (CLS) or partial least squares (PLS).[10]

Artificial neural networks (ANNs) are a set of emerging multivariate analysis tools with both qualitative and quantitative applications. These networks pass data from a multivariate input vector through one or more multivariate "layers" to an output layer. An illustration of an ANN is shown in Figure 1. The central layer of the network is termed a "hidden" layer because it is not evident from the standpoint of the input data or the results. Circles in Figure 1 represent data items or "nodes". The lines in the diagram represent the flow of data between layers and nodes. The input layer is identical with the input data vector of chemometric terminology, and the circles in the layer thus represent the elements of the input data vector. Circles in the hidden layer and output layer represent nodes or "neurons". A hidden layer node is formulated as shown in Equation 1. Here m is the number of input nodes, N_j is the value of node j, x_i is the value of input node i, w_i is the value of corresponding weight, b_i is the bias of the node, and f_j is known as the transfer function. Typical transfer functions used in classification problems include the log-sigmoid function, the tan-sigmoid function, or linear functions. These transfer functions are shown in Equations 2 – 4 respectively. The log-sigmoid function varies from 0 to 1 over a range of $-\infty$ to $+\infty$. The tan-sigmoid function varies from -1 to $+1$ over the range of $-\infty$ to $+\infty$. Nodes in the output layer can be formulated similarly, except x_i for an output layer node is the value of node i in the hidden layer.[11,12]

$$N_j = f_j(\sum_{i=1}^{m} w_i x_i + b_i) \qquad (1)$$

$$f(x) = \frac{1}{1+e^{-x}} \qquad (2)$$

$$f(x) = \frac{2}{1 - e^{-2x}} - 1 \qquad (3)$$

$$f(x) = x \qquad (4)$$

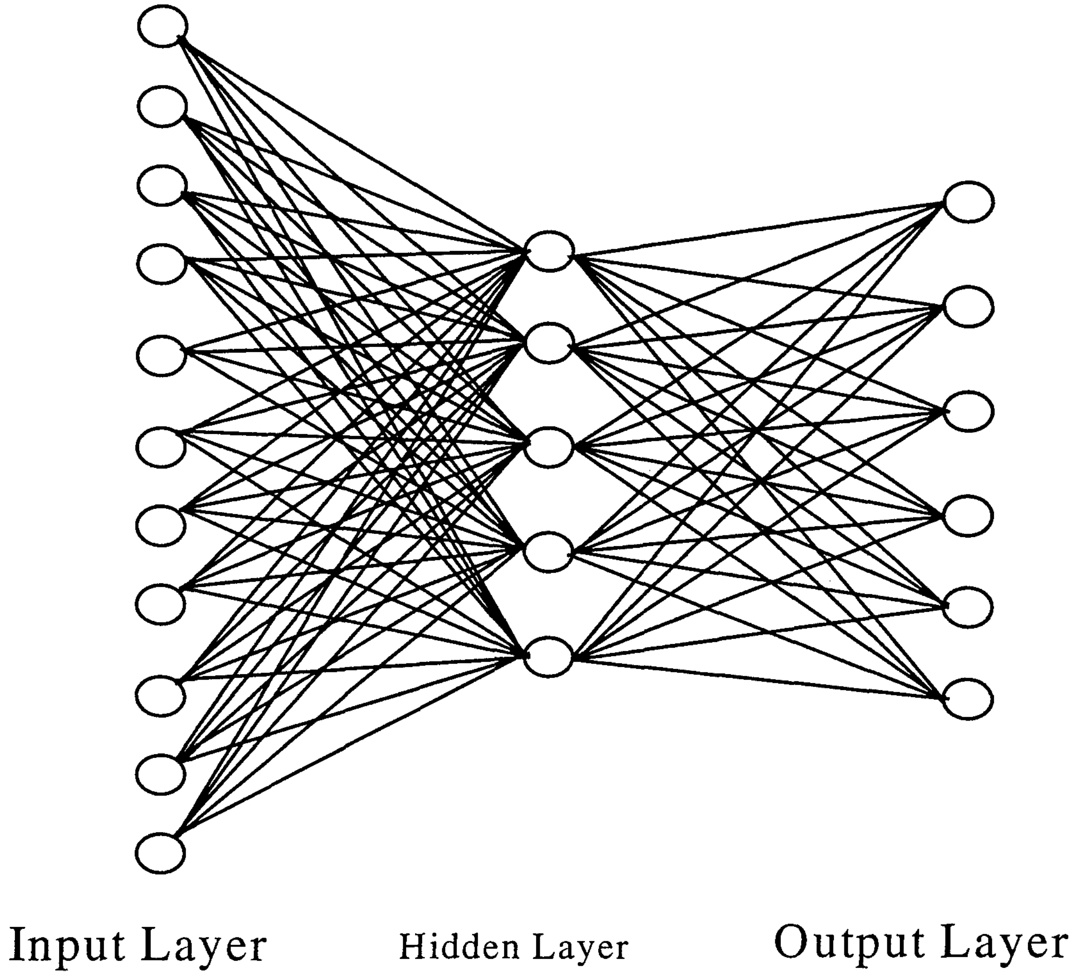


Figure 1. Diagram of an ANN. Circles represent nodes or neurons and lines represent data paths.

For classification problems, it is customary to formulate the output layer with a number of nodes equal to the number of classes under consideration. Each class is assigned a node in the output layer, and a data vector in a given class is represented with the value of 1.0 in the appropriate output node.[13] The challenge in using ANNs lies in setting the values of the weights and biases. Various optimization strategies can be applied to determining the optimum set of weights and biases for an ANN to ensure proper results. A commonly used general strategy is termed "back propagation". Here an initial set of weights and biases, usually set at random, are placed in the network and the network is presented with a set of input data and their associated correct classifications. Error signals obtained by comparison of the output layer with the correct output values are propagated back through the network and are used to adapt the weights and biases in order to improve the network performance. The process continues in an iterative fashion with the intention of improving the performance with each pass through the data set, hence the term "training". The details of this process give rise to a variety of training techniques.[11,12]

Neural network calculations for this work were carried out with the MATLAB® (Mathworks, Inc.) matrix mathematics package, supplemented by the Neural Network Toolbox. Several ANN training techniques were provided in this package, as well as routines for setting up the networks, initializing them, and evaluating them. Three training techniques selected include the gradient descent algorithm, the robust backpropagation algorithm, and the Levenberg-Marquardt algorithm. These techniques vary in speed of optimization, ability to determine a global optimum set of weights versus the tendency to be trapped in a local error minimum, and their use of computer memory. The gradient descent algorithm (GDA) has been used in a number of past chemometric evaluations[13] but it suffers from slow optimization and a tendency to be trapped by local error minima. Additional details have been given by Rumelhart, *et al.*.[14] The robust backpropagation technique (RPROP) usually exhibits improved optimization speed and less tendency to be trapped by local error minima. Additional details on the RPROP algorithm have been given by Reidmiller and Braun.[15] The Levenberg-Marquardt algorithm is based on the optimization techniques of the same name and utilizes the Jacobian matrix to establish weight and bias corrections. Additional details on the Levenberg-Marquardt algorithm have been given by Hagan and Menhaj.[16]

Radial basis function (RBF) networks are conceptually similar to ANNs but there are some differences in the formulation of nodes and in the method of training the networks. The schematic given in Figure 1 is still valid for RBF networks. Each node is formulated as shown in Equation 5, where N_j is node j, R is a radial basis transfer function, $\left\| x - w_j \right\|$ is the Euclidean vector norm of the difference between the vector of input nodes $\boldsymbol{x}$ and a weight vector $\boldsymbol{w}_j$, and b_j is the bias for the node. The transfer function typically is a Gaussian function, as shown in Equation 6, and it varies from 0 to 1 over a range of $+\infty$ to $-\infty$, with a maximum value of 1 at x=0.[12,17]

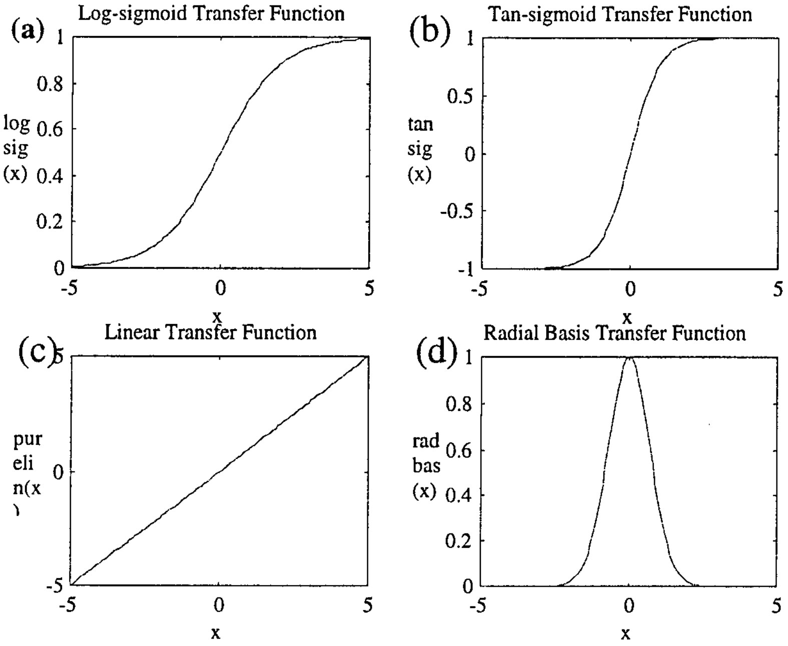
$$N_j = R(\left\| x - w_j \right\| \times b_j) \tag{5}$$

$$R(x) = e^{-x^2} \tag{6}$$

Radial basis functions are built by evaluating the network as new data vectors are presented. A new hidden layer node, or "center" is added when the evaluation of a new data vector produces output that causes the mean square error (MSE) value of the network to exceed a preset error goal. The weights and biases of the new node are mathematically determined, so that the network performs within the error goal with the new data vector included. Radial basis functions are often used for curve fitting and interpolation applications, and the output nodes produce real number values. For classification problems, the output layer can be spanned with node values from 0 to 1 and the product nodes can be rounded to the nearest integer.

The transfer functions are plotted over a limited domain in Figure 2. The log-sigmoid transfer function is plotted in Figure 2.a. This function concentrates numerical information from the entire real-number domain into the region between zero and one. The tan-sigmoid transfer function is plotted in Figure 2.b. This function is similar in shape to the tan-sigmoid function but its generated values range between +1 and –1. The linear transfer function, plotted in Figure 2.c transfers numerical values directly. The radial basis transfer function, plotted in Figure 2.d is radically different and this causes radial basis function neural network functions to behave differently from feed-forward networks. The radial basis transfer function produces the mathematical analog of a neuron that responds only to data within a restricted range. Input values far from zero, i.e. the center of the data domain, produce small response values form the function, while large output values result only from inputs very close to zero. As a result of this, radial basis function neurons tend to respond only to input data very close to a specific input pattern.

Figure 2. Neural network transfer functions. (a) Log-sigmoid function. (b) Tan-sigmoid function. (c) Linear function. (d) Radial basis transfer function.



Experimental

No infrared spectra were measured for these experiments. Mid-infrared absorption spectra were obtained as digital computer files from a variety of sources. Sadtler, Inc. provided condensed phase spectra of 48 pesticides and commercial organophosphorus chemicals. The US Army, military contractors, and other government organizations provided condensed phase and vapor phase spectra of a number of banned neurotoxins, their precursors, their hydrolysis products, and some commonly used simulants. These spectra were delivered in a variety of digital file formats, including simple X,Y ASCII files, spectra in a specific US Army format, JCAMP files, and files in proprietary GRAMS® (Galactic, Inc) and Nicolet formats. Some additional spectra were used from the NIST Gas Phase Infrared Library, which had been obtained in its optional JCAMP format.[18] In this option, the library is delivered as two JCAMP-formatted files, representing two spectral collections, one made by a contractor of the US-EPA and the other collected by NIST. The spectra used were translated into a common ASCII format which stored the chemical name and limited information regarding its source, and then stored the spectral observation points as X,Y pairs representing wavenumber and absorbance.

For the classification experiments, class 1 was chosen to represent 67 spectra of the banned neurotoxins, their precursors, hydrolysis products, and simulants. All class 1 compounds contained phosphorus, as non-organophosphorus substances were deleted from the study. The simulants were substances such as dimethylmethylphosphonate (DMMP) and diisopropylmethylphosphonate (DIMP) which are commonly used to test and demonstrate chemical warfare detection devices under benign conditions. Class 2 was chosen to represent the 48 organophosphorus pesticide spectra. Class 3 was chosen to represent compounds not belonging to either class 1 or class 2. The training set spectra for class 3 were the first 100 spectra from the NIST library JCAMP file, JCAMP.EPA. The class 3 training set spectra all represented substances that contained no phosphorus.

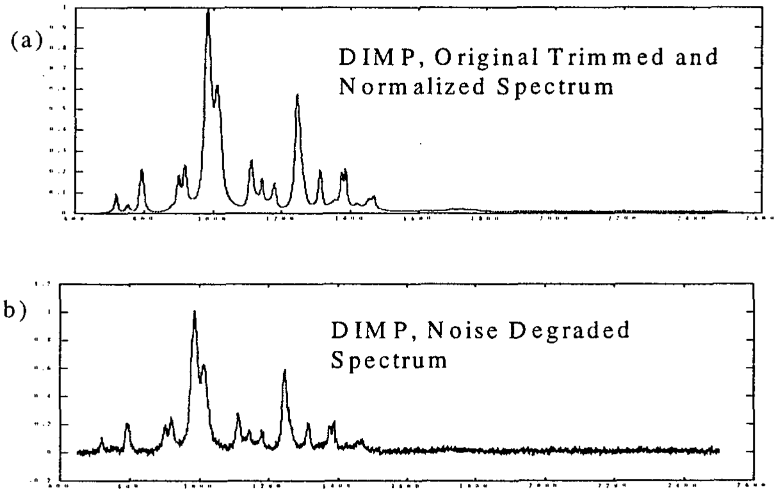
Initial classification experiments considered only the 115 spectra in classes 1 and 2. The available spectra were examined to determine the frequency range available in the spectra with the most restricted frequency range. A number of spectra in class 1 were found to have been stored in the frequency range from 650 to 2500 cm^{-1}, which was the most restricted frequency range. When preparing the spectra for the classification experiments, all spectra were "trimmed" by disregarding data outside this range. In order to further reduce artifacts relating the spectra to their various sources, it was found to be necessary to normalize the absorbance to the range of 0 to 1. The spectra thus prepared would not encode data necessary for quantitative estimates, but this loss did not interfere with the classification process. Each trimmed and normalized spectrum was transduced into a data vector by dividing the frequency range, from 650 to 2500 cm^{-1} into a number of equal width "bins" and averaging the absorbance values within each bin to generate the corresponding element of the data vector.

The number of spectra available for the class 1 and 2 compounds did not permit a subset to be removed and submitted as a test/evaluation set. Additional spectra could not be obtained beyond those used in the training set. Test/evaluation set spectra were partially synthesized by calculating a set of noise for each spectrum, using an infrared spectral noise model developed by Schuchardt.[19] The noise degraded spectra were trimmed, normalized, and transduced in the same manner as the training set spectra to produce test/evaluation data vectors for classes 1 and 2.

Feature selection was made by calculating Fisher weights for the data set in the manner described by Sharaf, et al.,[8] and removing those features whose weights were less than threshold values. Various thresholds were used to investigate the effect of the feature removal on the classification accuracy.

Chemometric calculations were carried out on a personal computer running the MATLAB matrix mathmatics package, (version 5.3) supplemented with the Neural Network Toolbox (version 3.0.1) and Statistics Toolbox (version 2.2) for MATLAB (The Mathworks, Inc.). The personal computer (Micron, Inc) was equipped with a Pentium-II® microprocessor and 64 Mbytes of RAM. Software ran under the Windows-NT operating system, version 4.0 (Microsoft, Inc.). Locally developed m-files controlled the MATLAB calculations.

Figure 3. DIMP Spectrum (a) before and (b) after being degraded with added noise.



Results

Figure 3.a shows the spectrum of DIMP, normalized to the absorbance range of 0 to 1, and trimmed to the frequency range of 650 – 2500 cm^{-1}. A noise-degraded spectrum produced by adding noise to the DIMP spectrum is shown in Figure 3.b. Two-class training and test/evaluation sets were generated by "binning" the 115 class-1 and class-2 original spectra and the 115 noise-degraded spectra into data four sets with 200, 100, 50, and 25 equal-width bins spanning the selected frequency range of 650-2500 cm^{-1}. The resulting bins had widths of 9.25 cm^{-1}, 18.5 cm^{-1}, 37 cm^{-1}, and 74 cm^{-1} respectively. The EPA vapor phase data set from the NIST library was also binned into a 25-bin data set over the same frequency range.

Initially, classification trials were made using the class 1 and class 2 spectra only, using data vectors binned from both the training set spectra and the noise-degraded test/evaluation set spectra. A principal component plot of this data set is shown in Figure 4. Trials using the *k*-nearest neighbor classifier (KNN) were not found to be sensitive to the voting committee size, *k*. Results using 3-nearest neighbors were typical, and they are summarized in Table 1. From the misclassifications detailed in Table 1, it appears that the optimum KNN classifications appear to come from the 100-bin data set, but the results are not overly sensitive to the bin-width, so that classifications from the 50-bin and 25-bin data sets are nearly as accurate. Calculating and applying Fisher weights produced a slight reduction in the numbers of misclassifications. Figure 5 shows a plot of the Fisher weight by bin number, as well as a plot of the average feature intensity.. The KNN results did not improve significantly when feature subsets were selected on the basis of their Fisher weights, although the errors did redistribute themselves between the classes.

Figure 4. Principal components plot of 100-bin, 2-class data set with Fisher weights applied.

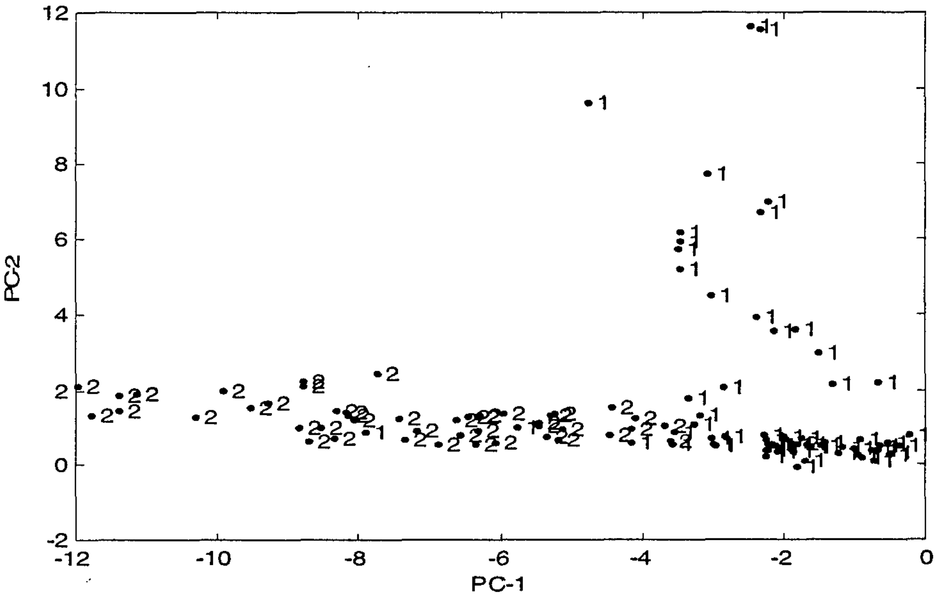
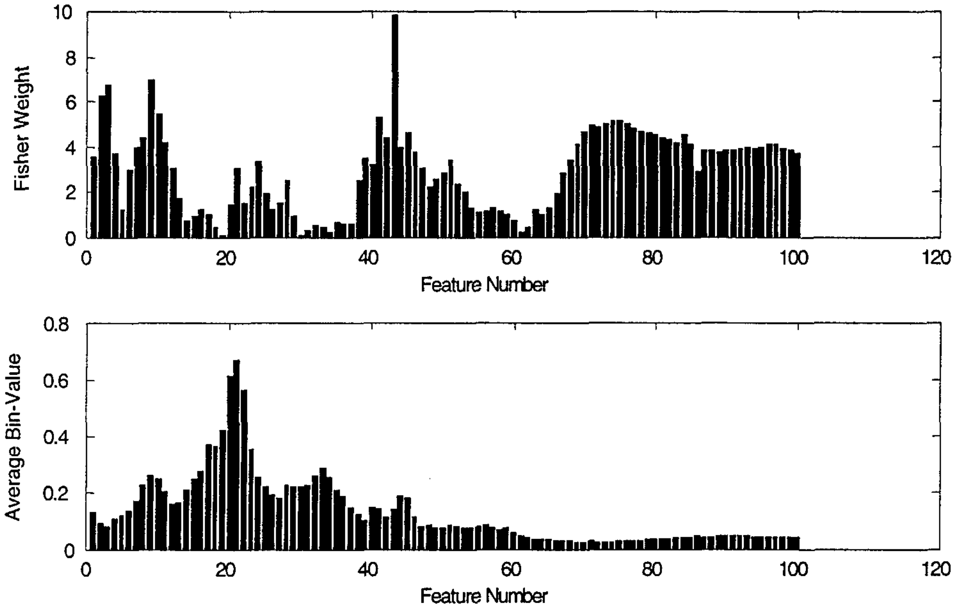


Table 1. Misclassification frequencies from 3-nearest neighbor classification, 2-class data set.

Data Set	3-NN, Raw Data	Fisher Weighted Data	Fisher Weights >3
200-bin	0.100	0.057	0.052
100-bin	0.083	0.057	0.061
50-bin	0.096	0.061	0.065
25-bin	0.070	0.061	0.065

Artificial neural networks were constructed and tested with routines in the MATLAB Neural Network Toolbox. ANNs to distinguish the class 1 and class 2 vectors were constructed with a single hidden layer, and a two-node output layer. Tan-sigmoid transfer functions were used for the input layer and hidden layer, and linear transfer functions were used for the output layer. Preliminary trials showed that five neurons in the hidden layer gave usable networks, and performance was similar using more or fewer hidden layer nodes. ANNs were trained to produce a set of integer output layer nodes with the pattern of [1 0] for class 1 data vectors and [0 1] for class 2 data vectors. The ANNs were trained to mean square error values of less than 10^{-3}, and for the final classification, the output layer nodes were rounded to the nearest integer value. The RPROP training routines in MATLAB proved to be the most generally useful of the three training systems investigated. ANNs trained with GDA training routines often failed to converge to the desired mean square error. The Levenberg-Marquardt training routines ran out of memory when data vectors featuring more than 30 bins were processed.

Figure 5. Plot of 2-class Fisher weights and feature average values in 2-class training set.



The MATLAB feed-forward ANN systems initialized each network with a unique set of weights and biases, and then proceeded to train the network. This led to some instability in the network evaluation and in the classifications obtained. If the network was successfully trained to a reasonable mean square error, it would usually identify all training set vectors without misclassification. There could be misclassifications in the test/evaluation set. It was found that setting up a committee of ANNs and averaging their output layers, the stability of the classifications could be improved, and the number of misclassifications could be reduced. Table 2 lists misclassification frequencies obtained from the 2-class data sets with RPROP-trained ANNs.

Figure 6. Principal component plot of 3-class training set, raw data.

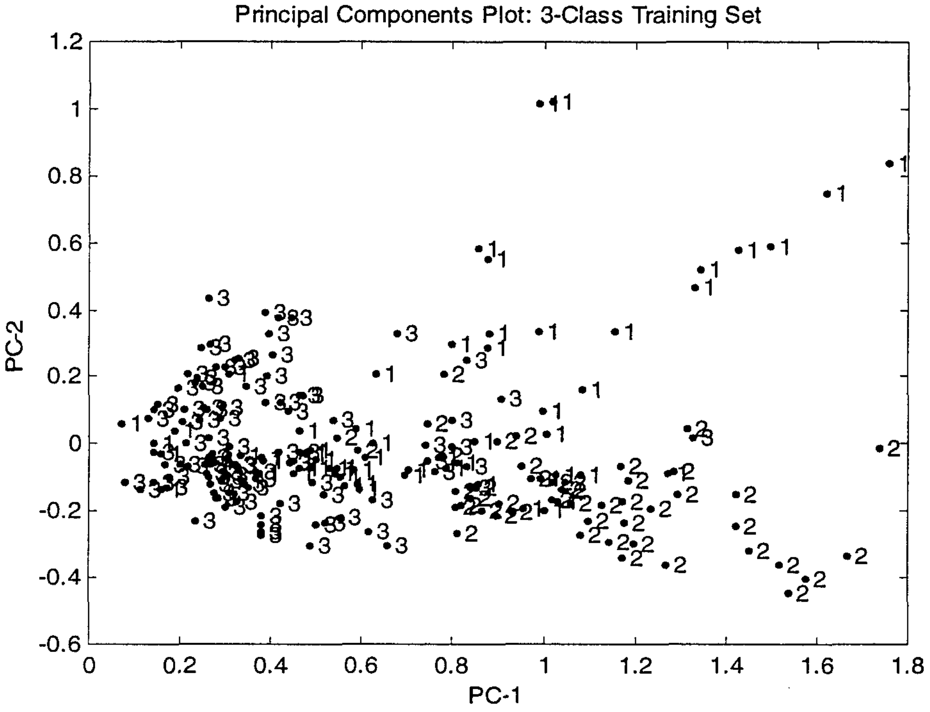


Figure 7. Principal components plot of 3-class training set with Fisher weights applied.

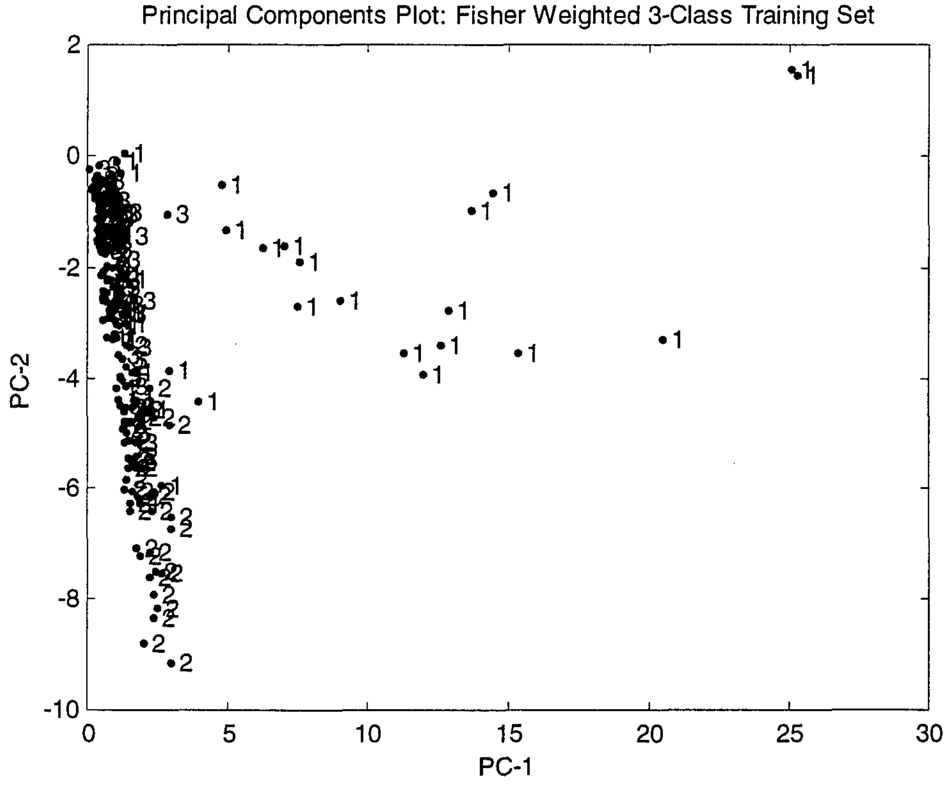


Table 2. Misclassification frequencies of selected ANN classifiers.

Data Set	Single Network Classifiers, Raw Data	5-Network Committee Classifiers, Raw Data	5-Network Committee Classifiers, Fisher Weighted Data	5-Network Committee Classifiers, Fisher Weights ≥ 3.0
200-bin	0.140	0.007	0.005	0.070
100-bin	0.005	0.014	0.003	0.035
50-bin	0.021	0.003	0.005	0.056
25-bin	0.520	0.012	0.012	0.050

Radial basis function networks have been constructed to process the 2-class data set. The results of the radial basis function network classifications have been given in more detail in a separate proceeding.

A training set was constructed from the 25-bin, 3-class spectral set. A principal component plot of this raw data set is shown in Figure 6. A principal component plot of the Fisher weighted data set is shown in Figure 7. KNN results from the raw training set indicate 22 total misclassifications from the raw 3-class training set. These misclassifications include 3 misclassifications of class 1 compounds as class2, and three misclassifications of class 1 compounds as class 3. These are potentially serious failures since warnings are particularly necessary from class 1 compounds. Additionally, there were 6 misclassifications of class 2 compounds as class 1, in effect false alarms, plus 2 misclassifications of class 2 compounds as class 3. There were 7 misclassifications of class 3 vectors as class 1, in effect false alarms again, and one false classification of a class 3 compound as class 2. Neural network classification testing of the 3-class data set has not been completed. ANN committees trained by the RPROP method classify the 3-class training set without error.

Discussion

This study required the collection of spectra from a diversity of sources. Prior to their use the spectra had to be translated into common digital formats to permit their mutual use. A number of the spectra obtained exhibited serious artifacts produced during collection, and these artifacts had to be removed or corrected prior to use. Many of the artifacts found in the spectra were outside the frequency range used here, so that they were eliminated when the spectra were trimmed. Other artifacts such as baseline height were corrected in the normalizing of the spectra. No duplicate pesticide spectra were available, and very few duplicates were obtained for the nerve agent, precursor, and hydrolysis product spectra. Thus, testing the classifications with authentic duplicate spectra was not practical.

The *k*-nearest neighbor algorithm proved to be simple to implement and it provided a stable classifier to use while investigating the effects of feature selection and other preprocessing techniques. The KNN classifier was relatively insensitive to the width of the binning in the preparation of the data vectors, even when the bin width grew to 74 cm^{-1}. This is in agreement with the finding by Griffiths that gas phase Fourier transform infrared spectra remained useful for quantitation and deconvolution procedures even when taken with resolutions as broad as 50 cm^{-1}.[21] Other workers have also reported that useful information remains in spectra acquired with resolutions as low as 16 cm^{-1}.[22]

Figure 5a indicates spectral regions with high Fisher weights between classes 1 and 2, and thus the spectral regions which are likely to be of use in descriminating between organophosphorus pesticides and the neurotoxins banned by the CWC, or their related precursors and hydrolysis products. Figure 5b shows the pattern of intensities within classes 1 and 2. Comparison of the two plots indicates that the most valuable spectral regions for the purposes of discrimination may not be the regions providing the most sensitive limits of detection for the compounds.

The ANN classifiers, trained with the RPROP technique appear to be a useful improvement over the KNN classifier from the standpoint of accurate discriminations. However, the numerical instability introduced by the random initialization of the network is troublesome. The use of a committee ANN approach improves the accuracy of the ANN further and it partially stabilizes the classifications.

Acknowledgement

This work was supported as an in-house effort by Air Force Institute of Technology, Department of Engineering Physics and by Air Force Research Laboratory, Airbase and Environmental Technology Division. This work was funded by the US Army, ERDEC. We also thank US Army ERDEC and Bio-Rad Laboratories, Inc., Sadtler Division for providing infrared spectra used in this work. We are also grateful to other organizations and workers who provided infrared spectra or technical assistance but asked to remain unacknowledged.

Disclaimer

Certain instruments and software have been identified by brand name to fully document the work. Such mention does not imply recommendation or endoresement by the Air Force nor does it imply that the items identified are the best available for the purpose.

References

1. P. R. Griffiths and J. A. de Haseth, *Fourier Transform Infrared Spectrometry*, John Wiley & Sons, New York, NY, 1986.

2. H. H. Willard, L. L. Merritt, Jr., J. A. Dean, and F. A. Settle, Jr., *Instrumental Methods of Analysis*, 7th Ed., Wadsworth Publishing Company, Belmont, CA, 1988.

3. J. Coates and J. Reffner, "Have FT-IR...Will Travel", *Spectroscopy*, **15**, 19 (2000).

4. CONVENTION ON THE PROHIBITION OF THE DEVELOPMENT, PRODUCTION, STOCKPILING AND USE OF CHEMICAL WEAPONS AND ON THEIR DESTRUCTION, (Corrected version in accordance with Depositary Notification C.N.246.1994.TREATIES-5 and the corresponding Procès-Verbal of Rectification of the Original of the Convention, issued on 8 August 1994).

5. G. D. Schuchardt, *Automated Infrared Detection of Organophosphorus Compounds in Multicomponent Solutions*, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1995.

6. R. S. McDonald and P. A. Wilks, Jr., *Applied Spectroscopy*, **42**, 151 (1988).

7. P. C. Jurs and T. L. Isenhour; *Chemical Applications of Pattern Recognition*, John Wiley and Sons, New York, NY, 1975.

8. M. A. Sharaf, D. L. Illman, and B. R. Kowalski, *Chemometrics*, John Wiley & Sons, New York, NY, 1986.

9. K. R. Beebe, R. J. Pell, and M. B. Seasholtz; *Chemometrics: A Practical Guide*, John Wiley & Sons, New York, NY, 1998

10. R. Kramer; *Chemometric Techniques for Quantitative Analysis*, Marcel Dekker, Inc., New York, NY, 1998.

11. S. Haykin, *Neural Networks. A Comprehensive Foundation*, Macmillan College Publishing Company, New York, NY, 1994.

12. H. Demuth and M. Beale, *Neural Network Toolbox for Use with MATLAB. User's Guide Version 3*, The Mathworks, Inc., Natick, MA, 1998.

13. J. R. Long, H. T. Mayfield, M. V. Henley, and P. R. Kromann, *Analytical Chemistry,* **63**, 1256 (1991).

14. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, Nature, **323**, 533 (1986).

15. M. Riedmiller and H. Braun, "A direct adaptive method for faster backpropagation learning: The RPROP algorithm," *Proceedings of the IEEE International Conference on Neural Networks*, 586 (1993).

16. M. T. Hagan and M. B. Menhaj, IEEE Transactions on Neural Networks **5**, 989 (1994).

17. S. Chen, C. F. N. Cowan, and P. M. Grant, *IEEE Transactions on Neural Networks*, **2**, 302 (1991).

18. *NIST Standard Reference Database 35. NIST/EPA Gas Phase Infrared Library in "JCAMP-DX" Format*, US Department of Commerce, 1992.

19. G. D. Schuchardt, *Automated Infrared Detection of Organophosphorus Compounds in Multicomponent Solutions*, MS Thesis, Air Force Institute of Technology, Wright-Patterson AFB, OH, 1995.

20. H. T. Mayfield, D. Eastwood, and L. W. Burggraf, "Classification of Infrared Spectra of Organophosphorus Compounds with Artificial Neural Networks" in Pattern Recognition, Chemometrics, and Imaging for Optical Environmental Monitoring, K. J. Siddiqui, D. Eastwood, Editors, Proceedings of SPIE Vol. 3854, pp. 56-64 (1999).

21. P. R. Griffiths, "FT-IR Spectrometry at Low Resolution: How Low Can You Go?", *Proceedings of the 9th International Conference on Fourier Transform Spectroscopy*, SPIE Proceedings Series, Vol. 2089, J. E. Bertie and H. Weiser editors, p. 2 (1994).

22. A. S. Bangalore, J. C. Demirgian, A, S, Boparai, and G. W. Small, Applied Spectroscopy, **53**, 1382 (1999).